View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

MATHEMATICS

AN INEQUALITY RELATED TO THE LARGE SIEVE

BY

DOUGLAS HENSLEY

(Communicated by Prof. J. H. van Lint at the meeting of June 28, 1975)

Introduction

In one formulation the large sieve reads as follows [7], [8]: Let $\mathcal{N}$ be a set of Z integers in the interval $[M+1, M+N]$. For prime p let $\omega(p)$ denote the number of residue classes mod p that contain no element of $\mathcal{N}$. Then for $x \geqslant 1$, $Z \leqslant (N+x^2)/L$, where

$$L = \sum_{n \leqslant x} \mu^2(n) \prod_{p|n} \frac{\omega(p)}{p-\omega(p)}.$$

In this paper we shall be concerned with estimating L when $\omega(p)=k$ ($k>1$ an arbitrary fixed integer) for all primes p such that $k<p\leqslant y$, (y a real parameter) and $\omega(p)=0$ for all other primes. In particular we examine the case $k=2$, which is of interest because of its connection with "twin prime" problems and related questions.

Notation. $\sum f(n)\ \{\quad\}$ and $\sum f(n) \prod g(p)\ \{\quad\}, \{\quad\}$ denote sums and products over the sets following them. Always n denotes a positive integer, and p, q primes.

The case $k=1$ is well understood. Van Lint and Richert [8] show that

$$\sum \mu^2(n) \prod (p-1)^{-1}\ \{n: n \leqslant x,\ p \nmid n \text{ if } p \geqslant y\},$$

$\{p: p|n\} = d(u) \log y + 0(1)$, where $u = \log x/\log y$, $d(u)=u$ for $0 \leqslant u \leqslant 1$, and $(d(u)/u)' = d(u-1)/u^2$ for $u \geqslant 1$.

Their work is based on the result of De Bruijn [4] about a related sum. Let $p(n)$ be the largest prime factor of n, and let

$$\Psi(x,y) = \sum 1\{n: n \leqslant x,\ p(n) \leqslant y\}.$$

Then $\Psi(x,y) = x\varrho(u) + 0(xu^2R(y)) +$ other error terms, where u is as before, $\varrho(u)=1$ for $0<u\leqslant 1$, and $u\varrho'(u) = -\varrho(u-1)$ for $u>1$.

In this paper we give similar estimates (with some loss of accuracy in the error term) for $\sum \prod 2/(p-2)\ \{n \leqslant x: n \text{ odd}, p(n) \leqslant y\}$, $\{p: p|n\}$, and for a sum bearing the same relation to the above as $\Psi(x,y)$ does to $\sum \mu^2(n)/\phi(n)\ \{n: n \leqslant x, p(n) < y\}$. We then generalize to $k>2$. The main result is at the end of § 1.

§ 1. $\Psi_2(x, y) \sim x \log x\, \varrho_2(u)$ (both to be defined).

DEFINITION. $p(n)$ denotes the largest prime factor of n; $q(n)$ the smallest.

DEFINITION. $\Omega(n)$ denotes the number of prime factors of n counting multiplicity.

DEFINITION. $\Psi_2(x, y) = \sum 2^{\Omega(n)}$ $\{n \leqslant x,\ 2 < q(n) \leqslant p(n) \leqslant y\}$.

DEFINITION.

$$S_2(x, y) = \sum \mu^2(n) \prod 2/(p-2) \ \{n \leqslant x,\ 2 < q(n) < p(n) \leqslant y\},\ \{p : p|n\}.$$

REMARK. This is not the same S_2 as in [8].

We have the following recurrence relation for $\Psi_2(x, y)$, which is fundamental in all that follows.

1) For $h > 1$, $\Psi_2(x, y) = \Psi_2(x, y^h) - 2 \sum \Psi_2(x/p, p)$ $\{p : y < p \leqslant y^h\}$.

LEMMA 1. $\Psi_2(x, x) = C_1\, x \log x + 0(x)$, where

$$C_1 = \prod (1 + 1/p(p-2))\ \{p \text{ odd}\}. \quad \square$$

For proof see [1], where the author shows that

$$\Psi_2(x, x) = C_1\, x \log x + Bx + 0(x^{\log 2/\log 3}).$$

We cannot use this extra accuracy since other estimates have larger errors.

We now define and examine a function which approximates

$$\Psi_2(x, y)/x \log x.$$

2) Let $\varrho_2(u) = C_1$ for $0 < u \leqslant 1$, $\varrho_2(u) = 0$ for $u \leqslant 0$, and

$$-u^2 \varrho_2'(u) = 2(u-1)\varrho_2(u-1) \text{ for } u > 1.$$

Let $v_2(u) = u\varrho_2(u)$. Then

3) $x \log x\, \varrho_2(u) = x \log x\, \varrho_2(u/h) -$

$$-2 \int^{y^h} \frac{x}{s} \log\left(\frac{x}{s}\right) \varrho_2\left(\frac{\log x}{\log s} - 1\right) \frac{ds}{\log s} \text{ if } u = \frac{\log x}{\log y},\ h \geqslant 1.$$

The following consequence of 2) is important for application of ideas in [2] and [3].

4) $uv_2(u) = 2 \int_0^1 v(u-t)dt$. Let

$$v_{2c}(u) = \frac{1}{2\pi i} \int_W \exp\{-uz + 2 \int_0^z s^{-1}(e^s - 1)ds\}dz,$$

where W is the contour coming from $-\pi i+\infty$ to $-\pi i$, then to πi, and finally to $\pi i+\infty$. Then

5) $uv_{2c}(u)=2\int_0^1 v_{2c}(u-t)dt$ (the necessary properties of the Fourier transform hold for W as well as for $(-\infty, \infty)$). The main theorem of [2] now gives $v_2(u)Cv_{2c}(u)(1+0(u^{-\frac{1}{2}}))$ for some real C.

REMARK. The asymptotic behavior of v_{2c} is more easily investigated than that of v_2. See [2], [3].

LEMMA 2. For $h>1$ let

$$\Delta(x, y, h)= \sum \Psi_2\left(\frac{x}{p}, p\right)\{y<p\leqslant y_h\}-\int_y^{y^h} \Psi_2\left(\frac{x}{u}, u\right)\frac{du}{\log u}.$$

Then

6) $$|\Delta(x, y, h)|=0\left[x\log x\, R(y)\left(h-1+\frac{1}{\log y}\right)\right],$$

where $R(y)$ is any function satisfying $R(y)\downarrow 0$,

$$R(y)>\frac{\log y}{y},\quad |\pi(y)-\mathrm{li}(y)|<\frac{y}{\log y}R(y).$$

REMARK. Basically this is an application of the prime number theorem.

PROOF OF LEMMA 2. Let $\Delta_n(x, y, h)$ be n's contribution to Δ. (Δ is a sum over certain integers).

$$\Delta_n(x, y, h)=2^{\Omega(n)}\left[\sum_{p\in J}-\int_J\frac{du}{\log u}\right]$$

where

$$J=\left[P(n), \frac{x}{n}\right]\cap[y, y^h].$$

If $n\leqslant x/y^h$ then $|\Delta_n(x, y, h)|\leqslant 2^{\Omega(n)}(4y^hR(y)/\log y)$. If $x/y^h<n\leqslant x/y$,

$$|\Delta_n(x, y, h)|\leqslant 2^{\Omega(n)}\left(4\frac{x}{n}\frac{R(y)}{\log y}\right).$$

Summing by parts over n, we have

$$|\Delta(x, y, h)|\leqslant 4xR(y)\left[(h-1)\log x-\tfrac{1}{2}(h^2-1)\log y+\left(\frac{\log x}{\log y}-h\right)\right]=$$

$$=0\left[x\log x\, R(y)\left(h-1+\frac{1}{\log y}\right)\right]. \qquad \square$$

Now we come to the key argument. Let

$$D(x, y) = \Psi_2(x, y) - x \log x\, \varrho_2 \left(\frac{\log x}{\log y}\right).$$

Then

$$\text{7)} \qquad D(x, y) = D(x, y^h) - 2 \int_y^{y^h} D\left(\frac{x}{u}, u\right) \frac{du}{\log u} - \Delta(x, y, h).$$

Taking absolute values in 7) we have

$$\text{8)} \qquad |D(x, y)| \leqslant |D(x, y^h)| + 2 \int_y^{y^h} \left| D\left(\frac{x}{u}, u\right)\right| \frac{du}{\log u} + |\Delta(x, y, h)|.$$

We are now in a position to state and prove

THEOREM 1: For $x, y > 2$

$$\Psi_2(x, y) - x \log x\, \varrho_2 \left(\frac{\log x}{\log y}\right) = 0(x) \log x\, R(y) \left(\frac{\log x}{\log y}\right)^2 + 0\left(x\left(\frac{\log x}{\log y}\right)^2\right).$$

PROOF. The idea is that Ψ_2 and $x \log x\, \varrho_2$ are close for $y = x$, and both satisfy the same recurrence relation (at least approximately) so as y decreases they will not separate much. For $y \geqslant x$, $|D(x, y)| = 0(x)$ by Lemma 1. Thus there exists $c > 0$ such that $|D(x, x)| < cx$ for $x \geqslant 1$. By Lemma 2 there also exists $c' > 0$ such that

$$|\Delta(x, y, h)| \leqslant c' x \log x\, R(y) \left[h - 1 + \frac{1}{\log y}\right]$$

for $x \geqslant 1$, $h > 1$, $y > 2$. Suppose that $|D(x, y)| \leqslant c_k x + c_k' x \log x\, R(y)$ for $y^k > x$. If $y^{k+1} > x$ and $h = 1 + 1/k$ then $(y^h)^k > x$. Thus

$$|D(x, y)| \leqslant c_k x + c_k' x \log x\, R(y) + 2 \int_y^{y^h} \left(c_k \frac{x}{u} + c_k' \frac{x}{u} \log \left(\frac{x}{u}\right) R(y)\right) \frac{du}{\log u} +$$
$$+ c' x \log x\, R(y) \left(\frac{1}{k} + \frac{1}{\log y}\right).$$

This is

$$\leqslant c_k \left(1 + 2 \log \frac{k+1}{k}\right) x + c_k' x \log x\, R(y) \left(1 + \frac{2}{k}\right) +$$
$$c' x \log x\, R(y) \left(\frac{1}{k} + \frac{1}{\log y}\right).$$

Therefore we take

$$c_{k+1}=\left(1+2\log\frac{k+1}{k}\right)c_k$$

and

$$c'_{k+1}=\left(1+\frac{2}{k}\right)+c'_k\left(\frac{1}{k}+\frac{1}{\log y}\right)c'.$$

One readily verifies that c_k and $c_k'=0(k^2)$. Since $y^k>x$ if and only if

$$\frac{\log x}{\log y}<k,$$

this completes the proof. □

§ 2. An asymptotic estimate for $S_2(x, y)$.

Recall that

$$S_2(x, y)=\sum \mu^2(n)\prod 2/(p-2)\ \{n\colon 2<q(n)\leqslant p(n)<y,\ n\leqslant x\}\ \{p\colon p|n\}.$$

Let

$$K_2(x, y)=\sum 2^{\Omega(n)}/n\ \{n\colon 2<q(n)\leqslant p(n)\leqslant y,\ n\leqslant x\}$$

and let

$$L_2(x, y)=\sum 2^{\Omega(n)}/n\ \{n\colon 2<q(n)\leqslant p(n)<y,\ n\leqslant x\}$$

(Identical unless y prime).

Lemme 3. $K_2(x, y)-L_2(x, y)\leqslant 1$ for y sufficiently large, independent of x.

Proof.

$$\sum \frac{2^{\Omega(n)}}{n}\ \{n\colon 2<q(n),\ p(n)=y,\ n\leqslant x\}\leqslant 2\sum\frac{2^{\Omega(n)}}{ny}$$

$$\{n\colon n \text{ odd},\ p(n)\leqslant y,\ n\leqslant x/y\}\leqslant \frac{2}{y}\prod_{p\leqslant y}\left(1-\frac{2}{p}\right)^{-1}\leqslant 1 \text{ for } y$$

sufficiently large. □

Lemma 4. $S_2(x, y)\geqslant L_2(x, y)$.

Proof. Let $\alpha(n)=\prod_{p|n}p$. Then

$$S_2(x, y)=\sum 2^{\Omega(n)}/n\ \{n\colon 2<q(n)\leqslant p(n)<y,\ \alpha(n)\leqslant x\}$$

because

$$\frac{2}{p-2}=\sum_{k=1}^{\infty} 2^{\Omega(p^k)}/p^k$$

so

$$\prod_{p|n} \frac{2}{p-2} = \sum 2^{\Omega(m)}/m\{m: \alpha(m) = a(n)\}.$$

If $\mu^2(n)=1$ then $\alpha(n)=n$, and the identity follows. Replacing $\alpha(n)$ by n, we have $S_2(x, y) \geqslant \sum 2^{\Omega(n)}/n \ \{n: n \leqslant x,\ 2 < q(n) \leqslant p(n) < y\}$. □

LEMMA 5. Let p^+ be the prime following p. Then $S_2(x, p^+) \geqslant S_2(x, p)$.

PROOF.

$$S_2(x, p^+) - S_2(x, p) = \frac{2}{p-2} S_2\left(\frac{x}{p}, p\right) \geqslant \sum_{n=1}^{\infty} \left(\frac{2}{p}\right)^n S_2(x/p^n, p).$$

Hence

$$S_2(x, p^+) \geqslant \sum_{n=0}^{\infty} \left(\frac{2}{p}\right)^n S_2(x/p^n, p). \quad \square$$

Similarly

$$L_2(x, p^+) = \sum_{n=0}^{\infty} \left(\frac{2}{p}\right)^n L_2(x/p^n, p).$$

Let $\delta_2(x, y) = S_2(x, y) - L_2(x, y)$. Then $\delta_2(x, p^+) \geqslant \delta_2(x, p)$. Thus

$$0 \leqslant S_2(x, y) - L_2(x, y) \leqslant \delta_2(x, x^+) = \sum \mu^2(n) \prod 2/(p-2)$$
$$\{n: n \leqslant x,\ 2 < q(n)\}\{p: p|n\} - \sum 2^{\Omega(n)}/n \ \{n: n \leqslant x,\ 2 < q(n)\}.$$

The residues at 0 of the Dirichlet series for $S_2(x, x)$ and for $L_2(x, x)$ have equal principal terms, so $\delta(x, x) = 0(\log x)^{2-1}$, and

$$S_2(x, x) - C(\log x)^2 = 0(\log x),\ L_2(x, x) - C(\log x)^2 = 0(\log x)$$

for some $C > 0$ (same C for both.

Now $L_2(x, y) = \int_{1-}^{x+} t^{-1} d\Psi_2(t, y)$. Integrating by parts and applying Theorem 1, we have

THEOREM 2:

$$L_2(x, y) = \log^2 y \int_0^u v\varrho_2(v)dv + \log y \int_{u-1}^u \varrho_2(v)dv +$$
$$+ 0(u^3 \log y) + 0(u^4 R(y) \log^2 y) + 0(\log x\ R(y) u^2) + 0(u^2) + 0(1).$$

COROLLARY: $S_2(x, y) =$ the same as above $+ 0(\log x)$.

§ 3. Generalization of Theorems 1 and 2 to $k > 0$ real.

As no new ideas are involved we omit most proofs. Let

$$C_k = \left(\prod_{p \leqslant k} \left(1 - \frac{1}{p}\right)^k\right) \exp\left\{\sum_{p > k} k \log\left(1 - \frac{1}{p}\right) - \log\left(1 - \frac{k}{p}\right)\right\} / (k-1)!$$

Let $\Psi_k(x, y) = \sum k^{\Omega(n)}$ $\{n: n \leqslant x, k < q(n) \leqslant p(n) < y\}$. Then

Lemma 1′: $\Psi_k(x, x) = C_k x(\log x)^{k-1} + 0(x(\log x)^{k-2})$.

Proof. One applies Theorem 2 of [9] with

$$f(s, k) = (\zeta(s))^{-k} \sum k^{\Omega(n)} n^{-s} \quad \{n: k < q(n)\}.$$

The "0" is uniform over any interval $[a, b]$ such that $(a, b]$ does not contain a prime. □

The recurrence relation for Ψ_k is

$$1)' \qquad \Psi_k(x, y) = \Psi_k(x, y^h) - k \sum_{y < p \leqslant y^h} \Psi_k(x/p, p).$$

In place of theorem 1, we have

Theorem 1′:

$$\Psi_k(x, y) = C_k\, x(\log x)^{k-1}\, \varrho_k(u) + 0(x(\log x)^{k-1}\, R(y) u^k) + 0(x(\log x)^{k-2} u^k)$$

where $\varrho_k(u) = C_k$ for $0 < u \leqslant 1$ and $-u^k \varrho_k'(u) = k(u-1)^{k-1} \varrho_k(u-1)$ for $u > 1$.

Let $v_k(u) = u^{k-1} \varrho_k(u)$. Then $uv_k(u) = k \int_0^1 v_k(u-t)dt$. Let

$$v_{kc}(u) = \frac{1}{2\pi i} \int_W \exp\{-uz + k \int_\iota^z s^{-1}(e^s - 1)ds\} dz$$

with W the same as before. Again $v_{kc}(u) = C(1 + 0(u^{-\frac{1}{2}}))v_k(u)$ for some real C.

Let

$$S_k(x, y) = \sum \mu^2(n) \prod \frac{k}{p-k} \{n: k < q(n) \leqslant p(n) < y, n \leqslant x\} \{p: p|n\}.$$

Let

$$K_k(x, y) = \sum k^{\Omega(n)}/n \quad \{n: k < q(n) \leqslant p(n) < y\},\ \{n \leqslant x\}$$

and

$$L_k(x, y) = \sum k^{\Omega(n)}/n \quad \{n: k < q(n) \leqslant p(n) < y,\ n \leqslant x\}.$$

Then

$$S_k(x, y) - L_k(x, y) = 0(\log x)^{k-1} \text{ independent of } y,$$

and

$$L_k(x, y) \sim (\log y)^k \int_0^u v_k(t)dt,$$

with leading error term of $0(u^{2k} R(y)(\log y)^k)$.

Remark. The "0" estimates are again uniform over any $[a, b]$ such that $(a, b]$ does not contain a prime.

REFERENCES

1. Bateman, P. T., Proof of a conjecture of Grosswald, Duke Math. Jornal **25**, No. 1, 67–72 (1958).
2. Bruijn, N. G. de, On some Volterra integral equations of which all solutions are convergent, Proc. Kon. Ned. Akad. van Wetenschappen **53**, 13–21 (1950).
3. ———, The Asymptotic Behavior of a function occurring in the theory of primes, J. Ind. Math. Soc. **15**, 25–32 (1951).
4. ———, On the Number of Positive Integers $\leqslant x$ and Free of Prime Factors $> y$, Proc. Kon. Ned. Akad. van Wetenschappen **54**, 50–60 (1951).
5. Ingham, A. E., The Distribution of Prime Numbers, Hafner Publishing Co., New York (1971). reprint.
6. Montgomery, H. L., Topics in Multiplicative Number Theory, p. 25, Springer-Verlag, New York (1971).
7. ——— and R. C. Vaugham, Hilberts Inequality, J. London Math. Soc. (2), **8**, 73–82 (1974).
8. Van Lint, J. H. and H. E. Richert, Über die Summe $\sum_{n \leqslant x,\, p(n) < y} \mu^2(n)/\phi(n)$, Proc. Kon. Ned. Akad. van Wetenschappen **66**, 582–587 (1964).
9. Selberg, A., Note on a paper by L. G. Sathe, J. Indian Math. Soc. **00**, 83–87 (1957).